Relationship between Allergic Contact Dermatitis and Electrophilicity

Herbert S. Rosenkranz,[1] Gilles Klopman,[2] Ying Ping Zhang,[1] Cynthia Graham,[1] and Meryl H. Karol[1]

[1]Department of Environmental and Occupational Health, University of Pittsburgh, Pittsburgh, PA 15238 USA; [2]Department of Chemistry, Case Western Reserve University, Cleveland, OH 44106 USA

To evaluate the role of electrophilicity in the induction of allergic contact dermatitis (ACD) in humans, we compared the structure–activity relationship (SAR) model of ACD with those of electrophilic and nonelectrophilic subsets of chemicals in the ACD database. For these analyses, electrophilicity was defined as the potential of a chemical to induce mutations in *Salmonella*. It was found that electrophilicity accounted for approximately 30–40% of ACD-inducing ability, and the remainder was associated with nonelectrophilic structures. The identification of these moieties opens the possibility for studying their role in ACD. *Key words*: allergic contact dermatitis, electrophilicity, mutagenicity, SAR, structure–activity relationship. *Environ Health Perspect* 107:129–132 (1999). [Online 13 January 1999]
http://ehpnet1.niehs.nih.gov/docs/1999/107p129-132rosenkranz/abstract.html

One of the mechanistic paradigms associated with the disorder allergic cutaneous dermatitis (ACD) is that the agent (i.e., electrophile or proelectrophile) responsible for eliciting this response attacks a nucleophilic cellular target (*1*). Others suggest that it is the overall reactivity that determines the response (*2*). Electrophilicity is also associated with mutagenicity and carcinogenicity (*3,4*). Mutagenic (genotoxic) carcinogens are considered to pose a greater risk to humans than nongenotoxic ones (*5*). The relationship between ACD and other health risks has recently been addressed by Albert (*6*), who suggested a relationship between ACD and nongenotoxic carcinogens. The present study was undertaken to evaluate this hypothesis, to determine the extent to which electrophilicity contributes to ACD, and possibly to identify other, nonelectrophilic, mechanisms that also contribute to ACD. Our approach was based on our earlier observation that the structural basis of ACD in humans could be modeled successfully by an expert structure–activity relationship (SAR) system (*7–9*) yielding a predictivity of 84%. The approach we have taken here was to derive SAR submodels for electrophilic and nonelectrophilic chemicals in the ACD database and to compare features of the two submodels.

Methods

CASE/MULTICASE. The algorithms employed within CASE and MULTICASE (Multicase, Inc., Beachwood, OH) have been described previously (*10–12*). Briefly, a set of chemicals and their respective experimentally determined biological activities (quantitative or qualitative) were input into a database. Both programs provide a means to identify descriptors consisting of molecular fragments, ranging from 2 to 10 heavy atoms along with their associated hydrogens, which account for the biological activity of the compounds under study. The molecular fragments are generated as a result of breaking down each individual chemical structure within the database into its constituent parts. Each fragment is "labeled" with respect to its origin within active or inactive structures. Fragments of relevance are those that exhibit a statistically significant nonrandom distribution among the active and inactive classes of compounds. In addition to using molecular fragments, MULTICASE identifies relevant two-dimensional distances between atoms within a chemical structure.

At this point within the analysis, CASE uses the statistically significant fragments in order to classify compounds as active or inactive. Predictions for submitted compounds are expressed as percent probabilities (0–100%) of being active or inactive. In addition, CASE utilizes the molecular fragments of the total database to derive a "global" quantitative structure–activity relationship (QSAR), which is used to predict the potency or extent of activity. These two predictions are derived independently of one another.

MULTICASE, on the other hand, uses the set of statistically significant descriptors (fragment and/or distance) to find a descriptor (biophore) that has the highest probability of being responsible for the observed biological activity. Compounds within the database that contain the primary biophore are removed from the analysis, and subsequent biophores that explain the activity of the remaining compounds are selected. This iterative process of selection is continued until either all of the active compounds are accounted for or no statistically significant descriptors remain. The presence of biophores determines the likelihood of a compound to exhibit activity. Predictions made on compounds submitted for SAR analysis consist of the identification of the biophore(s) responsible for activity and the percent probability of the compound being biologically active due to such occurrences. A compound is presumed to be inactive if it contains no biophores.

In MULTICASE, attempts are also made to derive a "local" QSAR within each group of compounds containing a particular biophore in order to identify molecular features that control the degree of activity. These features, termed modulators, are selected from the pool of molecular fragments, distance descriptors, calculated electronic indices (molecular orbital energies, charge densities), and calculated transport parameters (octanol/water partition coefficient, water solubility). These local QSARs are used to predict the potency of chemicals containing the specific biophore.

The similarity between structural determinants (i.e., biophores) associated with specific SAR models is taken to be a measure of mechanistic similarity (*13–15*). For this purpose, the biophores identified by the CASE program were used. The biophores selected were either identical (e.g., CH_2-CH_2- vs. CH_2-CH_2-) or embedded (e.g., CH_2-CH_2- vs. CH_2-CH_2-CH_2-). The structural overlap was defined as number of overlapping fragments/total number of significant fragments.

The database of mutagenicity in *Salmonella* was developed under the aegis of the National Toxicology Program (NTP) (*16–21*). An SAR model based upon subsets of that database has been described (*22*).

Database. The database used for SAR modeling was described previously (*7,8*). With respect to the present analyses in the model of ACD in humans, an "active" chemical elicits a response in human patch testing or has a sensitization rate of at least 16% in human maximization tests, when given at a dose greater than 10%. A chemical with "marginal" activity has a sensitization

Address correspondence to H.S. Rosenkranz, Department of Environmental and Occupational Health, University of Pittsburgh, 260 Kappa Drive, Pittsburgh, PA 15238-2818 USA.
MULTICASE, Inc. was cofounded and is partially owned by Case Western Reserve University, G. Klopman, and H.S. Rosenkranz.
Received 27 April 1998; accepted 23 September 1998.

rate of 4–15% under the same conditions. An "inactive" chemical is one that has a sensitization rate of <4% in human maximization testing.

Results and Discussion

To classify the chemicals in the ACD database into electrophiles/proelectrophiles and nonelectrophiles, we chose their potential to induce mutations in *Salmonella* as a surrogate for electrophilicity. This is based upon the recognition 1) that mutagenicity in *Salmonella* derives from an electrophilic attack and 2) that the *Salmonella* mutagenicity protocol includes testing in the presence of an exogenous metabolic activation mixture, thereby making it possible to identify proelectrophiles. Accordingly, we used an SAR model of the induction of mutations in *Salmonella* to classify chemicals with known ACD-inducing activity as potential electrophiles and nonelectrophiles. The predictivity of the model for *Salmonella* mutagenicity assay is 81%; however, because its sensitivity and specificity are approximately the same, the chance of false positive and false negative predictions is equal.

Table 1. Origin and composition of the three structure–activity relationship models

Characterization	ACD (*n*)	ACD-NE (*n*)	ACD-EL (*n*)
Total database	767	652	179
ACD^+	355	266	88
ACD_m	29	26	3
ACD^-	383	366	88
Potential nonelectrophiles	652	652	52
Potential nonelectrophiles ACD^+	266	266	0
Potential nonelectrophiles ACD^-	366	366	52
Potential electrophiles	114	0	127
Potential electrophiles ACD^-	88	0	36
Potential electrophiles ACD^+	23	0	88

Abbreviations: ACD, allergic contact dermatitis; NE, nonelectrophiles; EL, elecrophiles.
ACD^+, ACD_m, and ACD^- refer to chemicals shown experimentally to induce, marginally induce, and not induce allergic contact dermatitis, respectively. Potential electrophiles and potential nonelectrophiles refer to chemicals predicted to be mutagens and nonmutagens, respectively.

Table 2. Major CASE biophores associated with the parent allergic contact dermatitis structure–activity relationship model

Fragment 1	2	3	4	5	6	7	8	9	10	Fragments Sub	(*n*)	Inactives	Marginals	Actives	*p*-Value	No.
N	$-CH_2-$										50	1	0	49	0.000	1
N	$-CH_3$										26	1	0	25	0.000	2
N	=C–										35	2	0	33	0.000	3
NH	–C–										5	0	0	5	0.031	4
NH	–CH–										10	0	0	10	0.001	5
NH	$-CH_2-$										26	1	1	24	0.000	6
NH	$-CH_3$										5	0	0	5	0.031	7
NH_2	$-CH_2-$										7	0	0	7	0.008	8
NH_2	–NH–										5	0	0	5	0.031	9
Epoxide	$-CH_2-$										5	0	0	5	0.031	10
OH	–C=										89	10	0	79	0.000	11
F	–C–										6	0	0	6	0.016	12
S	–C–										5	0	0	5	0.031	13
S	–C=										14	0	0	14	0.000	14
SH	$-CH_2-$										19	0	1	18	0.000	15
Cl	$-CH_2-$										8	0	0	8	0.004	16
CS	–N–										5	0	0	5	0.031	17
CS	–S–										5	0	0	5	0.031	18
CH[a]	–CO	–CH=									15	0	0	15	0.000	19
N	–C	=CH									19	0	0	19	0.000	20
NH	–C	=CH–									19	0	1	18	0.000	21
CO	–C	–C–									10	1	0	9	0.006	22
CH	=C–	CH=								<2-NO_2>[b]	5	0	0	5	0.031	23
CH_2	–CH	$-CH_2-$								<2-Epoxide>	5	0	0	5	0.031	24
CH_2	=C	–CH–								<2-CO>	10	0	5	5	0.031	25
C	=CH	–CH	=C–								141	36	3	102	0.000	26
CH	=N	–CH	=CH–								8	0	0	8	0.004	27
NH_2	–C	=CH	–C=								5	0	0	5	0.031	28
NH_2	–C	=CH	–CH=								27	4	0	23	0.000	29
O	–CO	–C	$=CH_2-$								21	1	5	15	0.000	30
Cl	–C	=CH	–C=								13	0	0	13	0.000	31
CH	=C	–CH	=C–							<2-O>	5	0	0	5	0.031	32
CH	=C	–C	=C	–C=							5	0	0	5	0.031	33
CH	=C	–C	=C	–CH=							25	1	0	24	0.000	34
CH	=CH	–CH	=CH	–C[c]=							36	10	0	26	0.004	35
CH[3]	–C	=CH	–C	=CH–							11	1	0	10	0.003	36
OH	$-CH_2$	–C	=CH	–CH=							5	0	0	5	0.031	37
CH	=CH	–CH	=C	–CH=						<4-CO>	18	4	0	14	0.010	38
CH_2	$-CH_2$	$-CH_2$	–C	$-CH_2-$						<4-CH_3>	5	0	0	5	0.031	39
CH_2	$-CH_2$	$-CH_2$	$-CH_2$	$-CH_2$	–C=						15	0	0	15	0.000	40

< > indicates a substitutent. In biophore 36, the second and fourth carbon from the left are shown unsubstituted; this means they must be substituted by an atom other than hydrogen. On the other hand, in biophore 38, the unsubstituted fourth carbon from the left can only be substituted by a CO moiety, whereas the second carbon from the left in biophore 38 is shown with a hydrogen attached. No other atom can be at that location.

[a]This moiety is attached by a double bond to an outside substituent.

[b]A nitro substituent is on the second carbon atom from the left.

[c]This carbon atom is common to two rings.

Overall, it should be noted that while 28.0% of the 384 ACD-inducing agents were predicted to be mutagens, only 6% of the 383 noninducing chemicals were predicted to be mutagens. This suggests that electrophilicity is associated with ACD ($p \leq 0.0001$; χ^2 test), but that it is only one of a number of factors because it cannot explain the basis of the allergenicity of the remaining 72% of ACD-inducing agents.

The ACD data available consisted of 917 chemicals tested in humans; however, only 767 of these were used to construct the SAR model. Recent studies have shown that a ratio of active to inactive molecules in the learning set of unity was optimal for developing SAR models (*15*). Because inactive chemicals exceeded active ones, we randomly deleted some of them (using a computerized random number generator). Our final model consisted of 355 chemicals that were active in the human patch test or in the human maximization test, 29 marginally active and 383 inactive chemicals (*8*). As a result, we had a surplus of ACD noninducing chemicals that were not used to construct the ACD SAR model. A total of 652 of the 767 chemicals in the ACD database (85.0%) were predicted to be *Salmonella* nonmutagens (nonelectrophiles). Of these, 266 (40.8%) were inducers of ACD, 26 were marginal inducers, and 366 (56.1%) were noninducers of ACD (database: ACD-NE) (Table 1).

One hundred fourteen of the chemicals in the ACD database (14.9%) were predicted to be *Salmonella* mutagens (electrophiles). Of these, 88 (77.2%) were ACD inducers, 3 were marginal inducers, and 23 (20.2%) were noninducers. To increase the number of potential electrophilic ACD noninducers in the SAR model, we screened the "excess" ACD noninducers (n =150) that had not been used to construct the original ACD SAR model. An additional 13 potential electrophilic non-ACD inducers were thus identified. To bring the ratio of ACD-active to ACD-inactive to

Table 3. Major CASE biophores associated with the allergic contact dermatitis electrophilic (ACD-EL) structure–activity relationship submodel

Fragment										Fragments						
1	2	3	4	5	6	7	8	9	10	Sub	(*n*)	Inactives	Marginals	Actives	*p*-Value	No.
N	–C=										10	0	0	10	0.001	1
N	$-CH_2-$										13	1	0	12	0.001	2
N	$-CH_3$										7	0	0	7	0.008	3
NH	=C–										12	1	0	11	0.002	4
NH	–C=										5	1	0	14	0.000	5
NH	$-CH_2-$										10	0	0	10	0.001	6
NH_2	–C=										20	1	0	19	0.000	7
Epoxide	$-CH_2-$										5	0	0	5	0.031	8
S	–C=										9	0	0	9	0.002	9
Cl	$-CH_2-$										7	0	0	7	0.008	10
CO	–N–										5	0	0	5	0.031	11
CH	=C	–C=								<2-OH>[a]	7	1	0	6	0.035	12
CH	=C	–CH=								$<2\text{-}NO_2>$	5	0	0	5	0.031	13
CH_2	–CH	$-CH_2-$								<2-Epoxide>	5	0	0	5	0.031	14
C	=CH	–C	=C–								17	1	0	16	0.000	15
C	=CH	–CH	=C–								37	6	0	31	0.000	16
CH	=C	–CH	=C–								15	1	0	14	0.000	17
NO_2	–C	=CH	–CH=								11	0	0	11	0.000	18

< > indicates a substituent.

[a]A hydroxyl group is on the second carbon from the left.

Table 4. Major CASE biophores associated with allergic contact dermatitis nonelectrophilic (ACD-NE) structure–activity submodel

Fragment										Fragments						
1	2	3	4	5	6	7	8	9	10	Sub	(*n*)	Inactives	Marginals	Actives	*p*-Value	No.
N	$-CH_2-$										37	0	0	37	0.000	1
N	$-CH_3$										19	1	0	18	0.000	2
N	=C–										23	1	0	22	0.000	3
NH	–C=										14	0	1	13	0.000	4
NH	–CH–										6	0	0	6	0.016	5
NH	$-CH_2-$										16	1	1	14	0.000	6
NH_2	$-CH_2-$										5	0	0	5	0.031	7
OH	–C=										81	10	0	71	0.000	8
S	–C=										5	0	0	5	0.031	9
SH	$-CH_2-$										16	0	1	15	0.000	10
CO	–NH–										17	0	1	16	0.000	11
N	–C	=CH–									10	0	0	10	0.001	12
Cl	–C	=CH–									24	1	0	23	0.000	13
CO	–C	–C–									6	0	0	6	0.016	14
CO	–CH	$=CH_2-$									9	0	0	9	0.002	15
CH	=C	–CH=								$<2\text{-}NH_2>$[a]	6	0	0	6	0.016	16
CH	=C	–C	=C–								21	0	0	21	0.000	17
CH	=N	–CH	=CH–								8	0	0	8	0.004	18
O	–CO	–C	$=CH_2-$								19	1	5	13	0.000	19
OH	$-CH_2$	–C	=CH	–CH=							5	0	0	5	0.031	20
CH	=CH	–CH	=C	–C=						$<4\text{-}CH_2>$	8	0	0	8	0.004	21

< > indicates a substituent.

[a]An amino group is on the second carbon from the left.

Table 5. Summary of structural overlaps among allergic contact dermatitis (ACD) models

SAR model	Compared to	Overlap
ACD (total)	ACD	100%
	ACD-EL	42%
	ACD-NE	73%
ACD-NE (nonelectrophiles)	ACD-NE	100%
	ACD	95%
	ACD-EL	45%
ACD-EL (electrophiles)	ACD-EL	100%
	ACD	100%
	ACD-NE	63%

SAR, structure–activity relationship. Structural overlaps were determined as described in "Methods"; they are based upon the biophores of Tables 2, 3, and 4 as well as expanded fragments derived from them.

unity, we included an additional 52 ACD noninducers randomly selected from the data set (Table 1). The inclusion of potentially nonelectrophilic non-ACD inducers should not affect the SAR analyses to be performed, as we were interested only in activating moieties (biophores). Thus, this database (ACD-EL) consisted of 88 ACD inducers, 3 marginal inducers, and 88 noninducers (Table 1). Using CASE/MULTICASE we developed SAR models for the ACD, ACD-EL, and ACD-NE databases.

The major biophores associated with the "parent" ACD SAR model (Table 2) are, as expected, also present in the putative electrophilic (ACD-EL, Table 3) and nonelectrophilic (ACD-NE, Table 4) SAR submodels. Indeed, this is further evidenced by the extent of the structural overlap (i.e., commonality of fragments) between ACD on one hand and ACD-EL (42%) and ACD-NE (73%) on the other (e.g., Fragment 5 in Table 3 is identical to Fragment 4 in Table 4, etc.) (Table 5). This clearly suggests that the induction of ACD is not solely due to electrophilicity. However, as expected, the ACD-EL SAR model clearly contains typically electrophilic biophores (Table 3, biophores 7, 8, 18; e.g., arylamines, epoxides, and *ortho*-substituted nitroarenes) that are absent from the ACD-NE SAR model (Table 4). Further indication that ACD-EL and ACD-NE models reflect common, nonelectrophilic mechanisms of ACD induction is the substantial structural overlap between ACD-NE and ACD-EL, 45–63% (Table 5). Moreover, the ACD-NE model is highly predictive of the chemicals in ACD-EL and vice versa (i.e., 85 and 78% respectively, Table 6).

The findings described in this paper indicate that while electrophilicity is associated with the potential for inducing ACD, it is probably only one of the mechanisms involved. Clearly, the nonelectrophilic biophores identified in the two SAR submodels provide clues of alternate mechanisms whereby ACD may be induced. They need to be investigated further with respect to their origin and possible mechanistic significance.

Table 6. Cross-predictivities of the two structure–activity relationship (SAR) submodels

SAR model	Chemicals being predicted[a]	Concordance	*p*-Value*
ACD-NE (n = 652)	ACD-EL (electrophiles) (n = 179)	85.0%	<0.00001
ACD-EL (n = 179)	ACD-NE (nonelectrophiles) (n = 652)	78.1%	<0.00001

ACD, allergic contact dermatitis.
[a]The ACD-NE SAR model was used to predict the ACD-inducing ability of the chemicals used to derive the ACD-EL model and vice versa.
*Statistical significance of the concordance between experimental and predicted results.

Based on an examination of the activity of chemicals in the NTP rodent carcinogenicity bioassay and as ACD-inducing agents, Albert (*6*) suggested that there was a component among nongenotoxic carcinogens that was shared by human ACD-inducing agents. He reasoned that the results of ACD testing and mutagenicity testing in *Salmonella* could be complementary and could be used to identify carcinogens. Our results indicate that ACD is driven, to a large part, by nonelectrophilic mechanisms. Our findings, however, do not address the question of whether activity as an ACD-inducing agent can complement mutagenicity in *Salmonella* in predicting potential carcinogens.

References and Notes

1. Ashby J, Basketter DA, Paton D, Kimber I. Structure–activity relationships in skin sensitization using the murine local lymph node assay. Toxicology 103:177–194 (1995).
2. Benezra C, Sigman CC, Bagheri D, Fraginals R, Maibach HI. Molecular aspects of allergic contact dermatitis. In: Textbook of Contact Dermatitis (Rycroft RJG, Menné T, Frosch PJ, Benezra C, eds). Berlin, Heidelberg:Springer-Verlag, 1992;105–132.
3. Miller JA, Miller EC. Ultimate chemical carcinogens as reactive mutagenic eletrophiles. In: Origins of Human Cancer (Hiatt HH, Watson JD, Winsten JA, eds). Cold Spring Harbor, NY:Cold Spring Harbor Laboratory, 1977;605–627.
4. Ashby J, Tennant RW. Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the U.S. NTP. Mutat Res 257:229–306 (1991).
5. Ashby J, Morrod RS. Detection of human carcinogens. Nature 352:185–186 (1991).
6. Albert RE. Allergic contact sensitizing chemicals as environmental carcinogens. Environ Health Perspect 105:940–948 (1997).
7. Graham C, Gealy R, Macina OT, Karol MH, Rosenkranz HS. QSAR for allergic contact dermatitis. Quant Struct Act Relat 15:224–229 (1996).
8. Graham C. Development and comparisons of structure–activity relationship models of chemical respiratory and dermal hypersensitivity [Ph.D. dissertation]. University of Pittsburgh, Graduate School of Public Health, Department of Environmental and Occupational Health, Pittsburgh, PA, 1998.
9. Johnson R, Macina OT, Graham C, Rosenkranz HS, Cass GR, Karol MH. Prioritizing testing of organic compounds detected as gas phase air pollutants: structure–activity study for human contact allergens. Environ Health Perspect 105:986–992 (1997).
10. Klopman G. Artificial intelligence approach to structure–activity studies. Computer automated structure evaluation of biological activity of organic molecules. J Am Chem Soc 106:7315–7321 (1984).
11. Klopman G. MULTICASE 1. A hierarchical computer automated structure evaluation program. Quant Struct-Act Relat 11:176–184 (1992).
12. Klopman G, Rosenkranz HS. Prediction of carcinogenicity/mutagenicity using MULTICASE. Mutat Res 305:33–46 (1994).
13. Rosenkranz HS, Klopman G. Structural evidence for a dichotomy in rodent carcinogenesis: involvement of genetic and cellular toxicity. Mutat Res 303:83–89 (1993).
14. Rosenkranz HS, Zhang YP, Klopman G. Evidence that cell toxicity may contribute to the genotoxic response. Regul Toxicol Pharmacol 19:176–182 (1994).
15. Liu M, Sussman N, Klopman G, Rosenkranz HS. Structure–activity and mechanistic relationships: the effect of chemical overlap on structural overlap in data bases of varying size and composition. Mutat Res 372:79–85 (1996).
16. Dunkel V, Zeiger E, Brusick D, McCoy EC, McGregor D, Mortelmans K, Rosenkranz HS, Simmon VF. Reproducibility of microbial mutagenicity assays. I: Tests with *Salmonella typhimurium* and *Escherichia coli* using a standard protocol. Environ Mutagen 6:1–251 (1984).
17. Dunkel VC, Zeiger E, Brusick D, McCoy E, McGregor D, Mortelmans K, Rosenkranz HS, Simmon VF. Reproducibility of microbial mutagenicity assays. II: Testing of carcinogens and noncarcinogens in *Salmonella typhimurium* and *Escherichia coli*. Environ Mutagen 7:1-248 (1985).
18. Mortelmans K, Haworth S, Lawlor T, Speck W, Tainer B, Zeiger E. *Salmonella* mutagenicity tests. II: Results from the testing of 270 chemicals. Environ Mutagen 8(suppl 7):1–119 (1986).
19. Zeiger E. Carcinogenicity of mutagens: predictive capability of the *Salmonella* mutagenesis assay for rodent carcinogenicity. Cancer Res 47:1287–1296 (1987).
20. Zeiger E, Anderson B, Haworth S, Lawlor T, Mortelmans K, Speck W. *Salmonella* mutagenicity tests. III: Results from the testing of 255 chemicals. Environ Mutagen 9:1–109 (1987).
21. Zeiger E, Anderson B, Haworth S, Lawlor T, Mortelmans K. *Salmonella* mutagenicity tests. IV: Results from the testing of 300 chemicals. Environ Mol Mutagen 11:1–157 (1988).
22. Rosenkranz HS, Klopman G. The structural basis of the mutagenicity of chemicals in *Salmonella typhimurium*: the National Toxicology Program data base. Mutat Res 228:51–80 (1990)